



# AI FOR CALORIMETRY/HACKATHON

---

William Phelps for Team Oxpecker Reborn

Christopher Newport University/Jefferson Lab

# What is a Hackathon?

- It can mean many things but, in this context, it is a competition and a learning experience with relevant problems usually generated w/toy models
- Specifically, we will talk about the AI hackathon that took place on July 17<sup>th</sup>, after the AI Town Hall
  - But you will see a few photos of the hackathon in 2020!
- Typically, there are teams with ~4-5 people
  - Not everyone has to be experienced
- Held over the course of one day with NP based problems/datasets



Hackathon in March 2020 – AI4NP

# Hackathons

- Foster community engagement
- Great learning experience
- **And most of all, fun!**

## Oxpecker Reborn Team Roster

- Gagik Gavalian (JLab)
- Tyler Viducic (ODU)
- Andru Quiroga (CNU)
- Torri Jeske (JLab)
- William Phelps (CNU/JLab)



Hackathon in March 2020 – AI4NP



AI Townhall Hackathon - July 2021

# Hackathon Logistics

- 6 Teams ~50 participants
- Website with challenge info and datasets
- Email address to automatically judge submissions!
- Run by the EPSCI and Data Science groups at Jefferson Lab

Thomas Britton  
Diana MacSpadden  
Kishan Rajput  
Chris Tennent

## Overall

| Team Name            | Overall Score   |
|----------------------|-----------------|
| Oxpecker Reborn      | 235.82656552869 |
| The Nervous Networks | 172.99484296064 |
| winners              | 88.545993151835 |
| Proton               | 87.168277431949 |
| Team Turtle          | 80.852772127346 |
| The Golden Griffin   | 27.83775220045  |

## Problem 1

| Team Name            | Problem Score      |
|----------------------|--------------------|
| Oxpecker Reborn      | 8.163788238695005  |
| The Nervous Networks | 66.86754251577614  |
| Proton               | 79.52365968464751  |
| Team Turtle          | 1918.1965562584846 |
| winners              | 7956.632784857568  |
| The Golden Griffin   | 40378.38170495783  |

## Problem 2

| Team Name            | Problem Score      |
|----------------------|--------------------|
| The Nervous Networks | 61.7931339272663   |
| Oxpecker Reborn      | 68.03510353504575  |
| Team Turtle          | 435.6676160483719  |
| The Golden Griffin   | 2935.1811899895774 |
| winners              | 3123.4748319773944 |

## Problem 3

| Team Name            | Problem Score      |
|----------------------|--------------------|
| Oxpecker Reborn      | 97                 |
| The Nervous Networks | 209                |
| Proton               | 381.04323380821165 |
| Team Turtle          | 641                |
| winners              | 736                |

## Problem 4

| Team Name            | Problem Score      |
|----------------------|--------------------|
| Oxpecker Reborn      | 39818.34017962084  |
| Proton               | 85387.24586274063  |
| The Nervous Networks | 90239.95624025587  |
| Team Turtle          | 110588.26762290485 |
| winners              | 111563.48776312581 |

## Problem 5

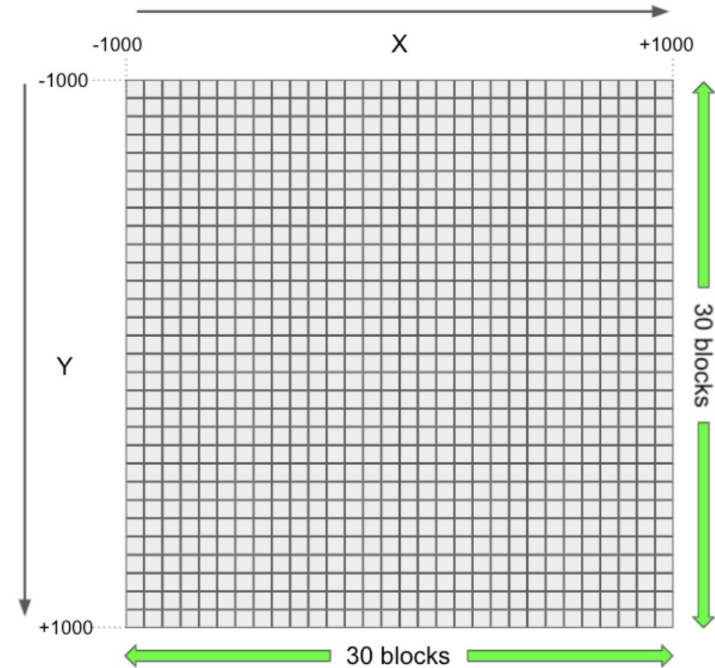
| Team Name            | Problem Score      |
|----------------------|--------------------|
| Oxpecker Reborn      | 104519.26073971845 |
| winners              | 186241.24325047343 |
| The Nervous Networks | 217902.8833570769  |

Live scoreboard

# Problem sets

- 30x30 simulated calorimeter
- 5 problems of increasing difficulty
- Started out with events with single hits and no noise
- Advanced problems contained multiple hits with noise
- You are given training data and labels in csv format

Note: These problems/solutions may not be ideal! We had a finite amount of time



# Tools of the Trade

- Python 3.8 – Anaconda
  - Keras/TensorFlow - NN Libraries
  - Pandas/Numpy - Data Handling
  - Matplotlib - Visualization
- Local Machines, cloud resources such as google Colab, or Jlab Scientific Computing resources
  - 4 Titan RTX cards per node
  - Soon there will be nodes with 16 Nvidia T4 cards available

```
test = pd.read_csv("TRAIN/TRAIN.csv")
labels = pd.read_csv("TRAIN/TRAIN_labels.csv")
activation = 'relu'

model = Sequential()
model.add(Dense(units=1000, activation=activation, input_shape=(3600, )))
model.add(Dense(units=1000, activation=activation))
model.add(Dense(units=1000, activation=activation))
model.add(Dense(units=2))
model.compile(optimizer=adam(lr=.001), loss='mean_squared_error', metrics=['accuracy'])

model.fit(test, labels[labels.columns[1:]], epochs=300, batch_size=256, validation_split=0.2)
```

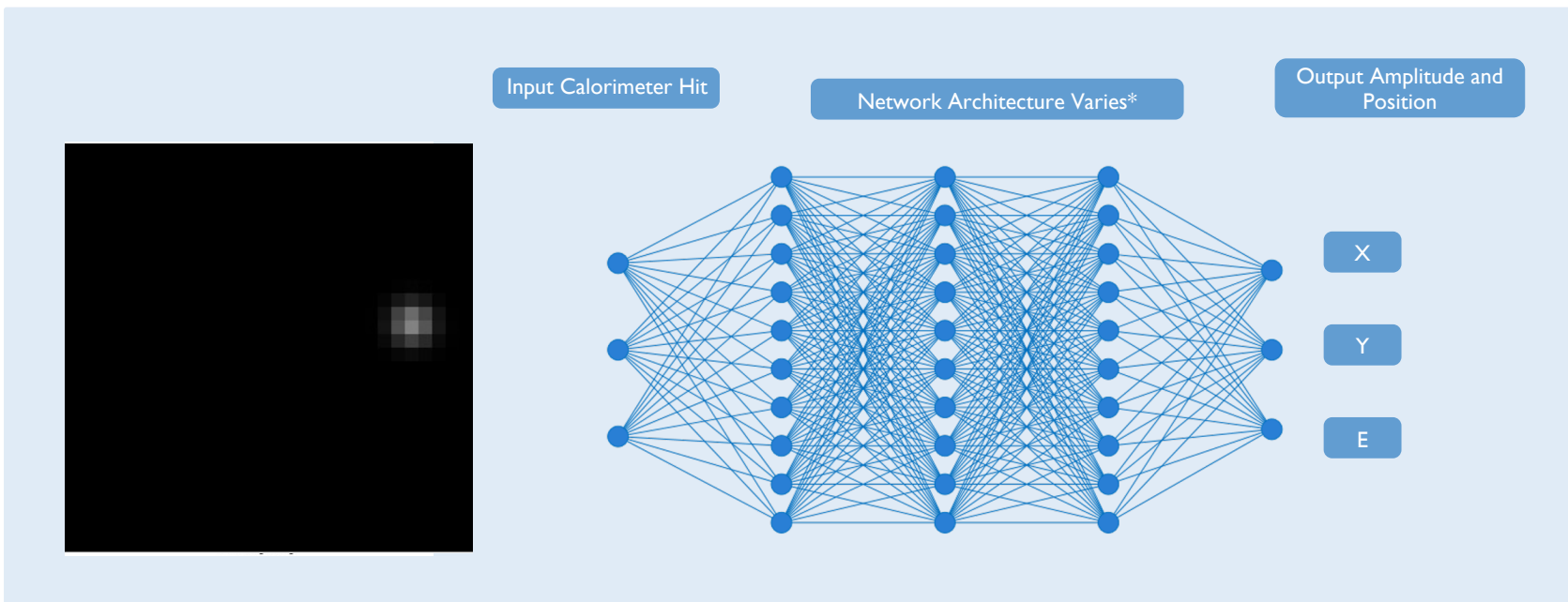
Sample Training Script



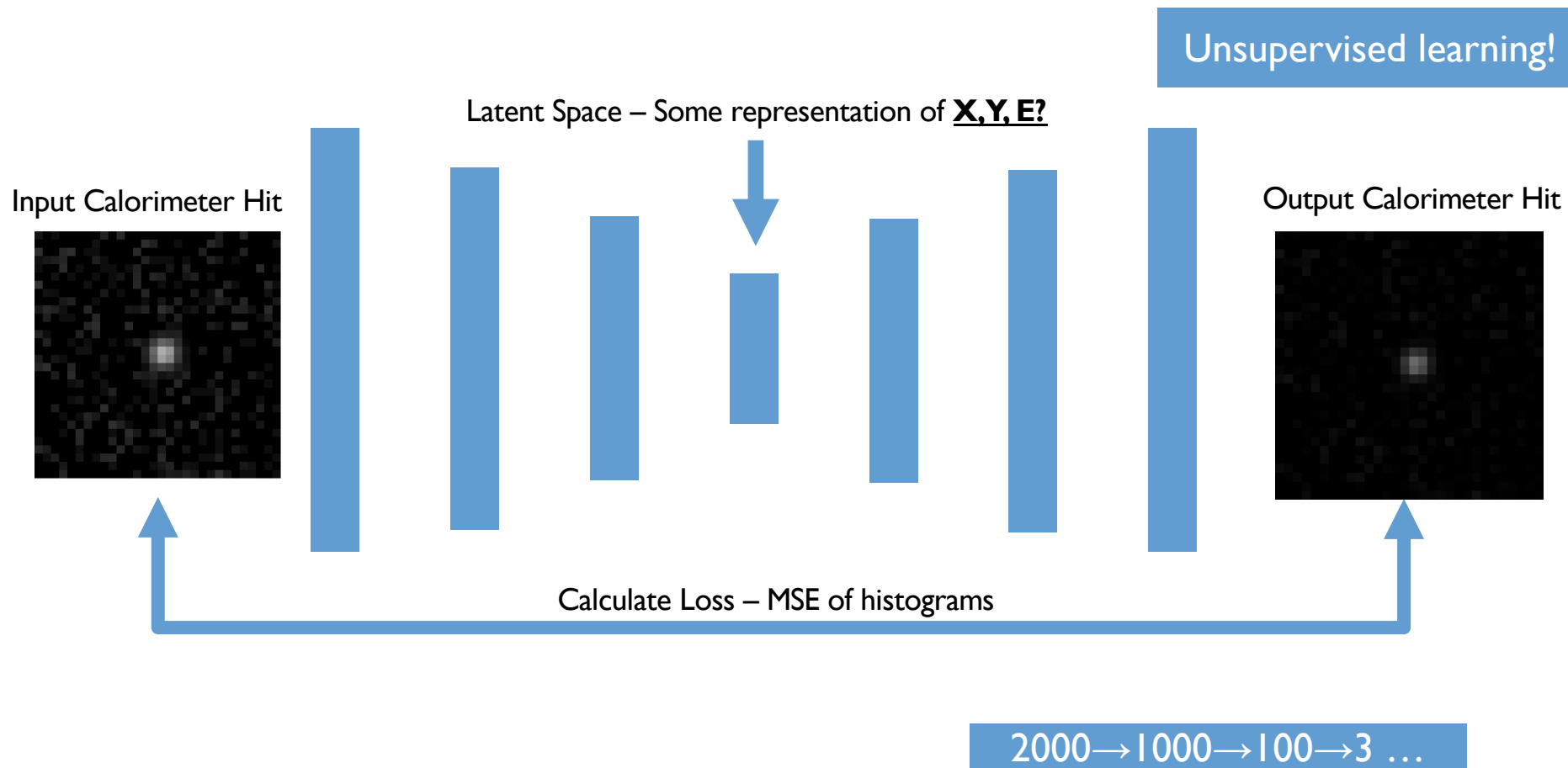


# Problem I – Single Hit

- Read in 30x30 pixel hits from CSV files, provide calorimeter hit position (x, y) and amplitude of the hit.
- Model Architecture: 3 layers of 2000 neurons, relu activation function



## Problem 2 – Denoising

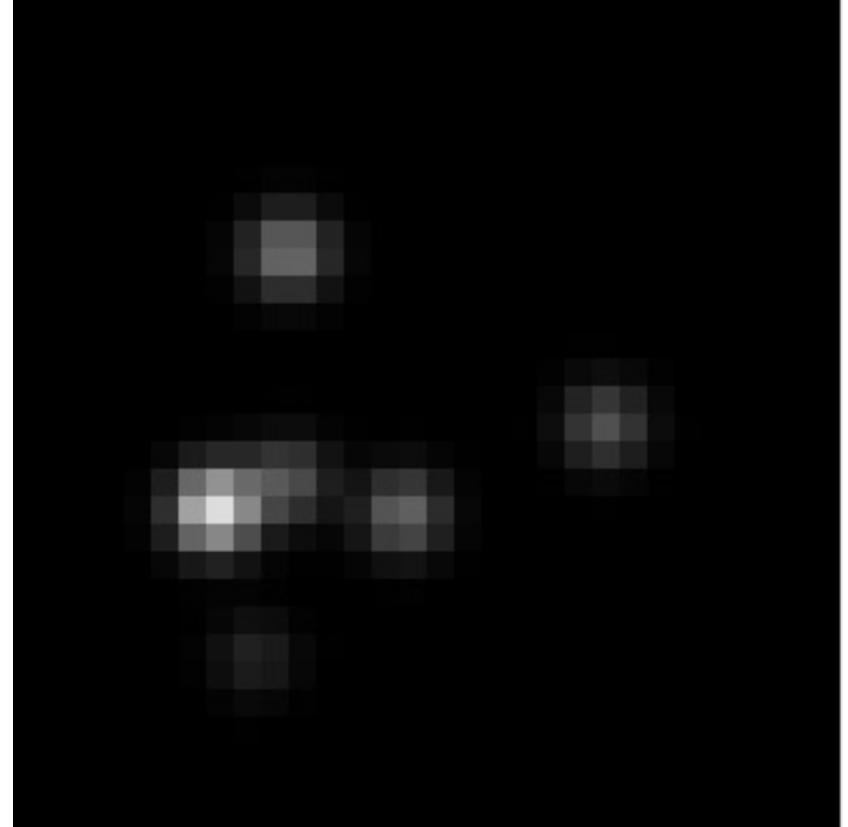




## Problem 3 – Count clusters

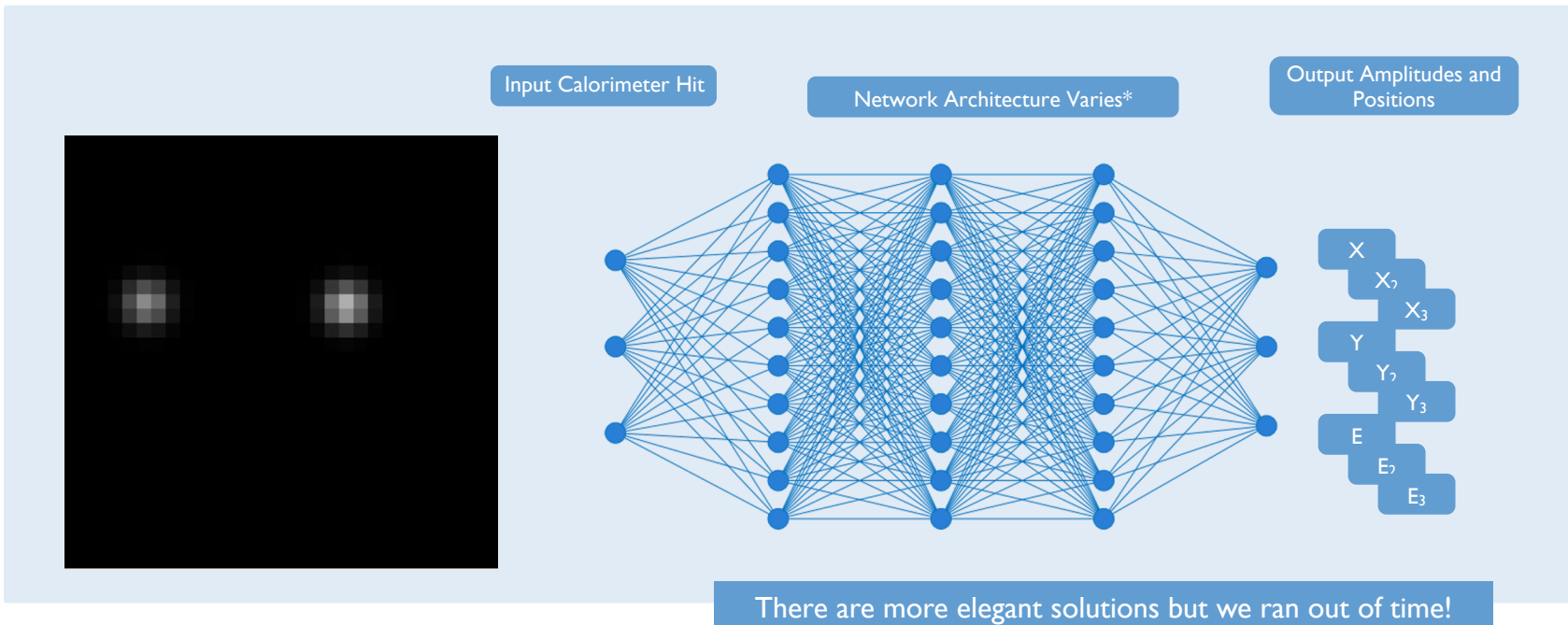
- Multiple showers per event from a Poisson distribution
- No upper limit on # showers
- Goal: Return the # of showers given a calorimeter hit (shown on the right)

Solved using a Convolutional Neural Network



# Problem 4, 5 – Multiple Hits, w/Noise

- Read in 30x30 pixel images from CSV files, provide calorimeter hit position(s) (x,y) and amplitude of the hit.
- Model Architecture: 5 layers of 2000 neurons, relu activation function



# Summary

- Hackathons are a great experience!
  - If you are paying attention and you thought these problems are interesting then you should sign up
- They make a great educational experience because typically we are all familiar with how to wrangle data but not necessarily the AI tools
- When someone else does the wrangling and presents a clear problem you can focus on the ML techniques
- Is the EIC community interested in having an AI Hackathon?